Human Beatboxing : A preliminary study on temporal reduction.

Alexis Dehais Underdown[1], Paul Vignes[1], Lise Crevier Buchman[1,2] & Didier Demolin[1]
[1] Laboratoire de Phonétique et Phonologie (Sorbonne Nouvelle / CNRS), [2] Hopital Foch, Univ. VSQ.
alexis.dehais-underdown@sorbonne-nouvelle.fr

1. Introduction

Speech rate is known to be a factor of reduction affecting supralaryngeal gestures [4] [6] [1] and laryngeal gestures [5] depending on the prosodic structure [3]. An acoustic study [2] showed that beatboxing rate has articulatory effects on hi-hats in medial and final positions but that the overall production was not found to be affected. In the present study we are presenting an experiment based on a beatboxing speeding up paradigm. We used a single metric and rhythmical pattern to create various beatboxed patterns (BP). Human Beatboxing (HBB) seems to rely on different articulatory skills than speech because it does not obey to linguistic constraints. How do beatboxing rate affect sound duration ? We expect that (1) the faster the production, the shorter sound duration (2) affricates, trills and fricatives will shorten more than stops (3) position in the pattern influences sound reduction.

2. Methods

A 32 y.o. professional beatboxer took part in the experiment. The recorded material was controlled in terms of sound pattern and rhythm. Indeed, to facilitate the comparison of beatboxing recordings we used only one metric pattern that we transposed into 12 Beatbox patterns (BP). Table 1 shows an example of two of the 12 BP created for the study. There is only one voiced sound (i.e. [ɓ]) in the corpus.

Table 1 : Example of the structure of a Beatboxed Pattern

Metric									
Rhythm	Low	High	Long	High	Low	Low	High	Long	High
Positions	1	2	3	4	5	6	7	8	9
BP1 **BP2**	[p’] [↓ʙ̥ˡ]	[t͡s’] [t͡s’]	[↓k͡L̥] [ʔhə̥]	[t͡s’] [t͡s’]	[p’] [↓p]	[p’] [↓ʙ̥ˡ]	[t͡s’] [t͡s’]	[↓k͡L̥] [ʔhə̥]	[t͡s’] [t͡s’]

Concerning the rhythm, it was controlled using a vibrating metronome placed on his wrist. The participant was asked to repeat each BP 5 times at 90 Beat per Minute (BPM), 120 BPM and 150 BPM (12 BP x 5 repetition x 3 speed = 180 BP).
To study temporal reduction, we extracted sound duration (ms), acoustic phase duration (ms) (e.g. frication, trilling), silence between sounds (ms) and for each pattern we ploted durations of sounds and speed rate.

3. Results

3.1. Sound Duration

Concerning sound reduction, global results show sounds shorten when speed rate is increasing. However, when we have a close look at the data we found that sounds reduce differently. For example [ɓ] lengthen rather than shortening. Similarly, we found that the guttural snare [↓ʔə̥h] (the arrow indicates an ingressive flow and the voiceless schwa indicates the timbre) shortens in a particular way : its glottal onset (i.e. ingressive voicing) lengthens but the aspiration noise shortens; though the overall duration shortens. This suggests that there may be differences between supralaryngeal control, laryngeal control and “ingressive phonation” that needs to be further understood. Finally, since there is no vowels, we measured the duration of silence between sounds. The data shows the silences shorten a lot. Thus, the temporal distance between gesture decreases as the speed rate increases; that is, the time interval to move from a target to another is shortened which can lead to articulatory overlap.

3.2. Mode of articulation

We hypothesized that affricates, trills and fricatives will shorten more than stops. This seems to be the case but all sounds do not reduce the same. The voiceless ingressive pulmonic trill [↓ʙ̥ˡ] is longer than its velic correlate {ŋ[↓ʙ̥ˡ]ŋ} (brackets indicate that humming is possible) and thus shortens more. The oral affricates (e.g. [↓k͡ʟ̥], [ǂ*], [t͡ʃi̥]) reduce the duration of the frication noise except for the labiodental affricates [pf'] (i.e. glottalic) and {ŋ[pf]ŋ} (i.e. velic) in third position. Given that they appear in the context in position 3 and 8 we do not have an explanation for their lengthening. The "guttural" affricate [↓ʔi̥ʃ] reduces very much its frication noise whereas [↓ʔə̥h] and [↑ʔɛ̥ʃ] shorten only a bit. Finally, the velic fricative {ŋ[f]ŋ} do not reduce as we thought, it only reduces in final position. Velic fricatives are not attested, to our knowledge, in the world's languages. Aerodynamic and articulatory data may help us understand the constraints on this type of sound that could allow us to interpret the results.

3.3. Position

The third hypothesis was that position (cf. table 1) influences sound reduction. The final position (i.e. 9) favors sound reduction since we noticed all hi-hats shorten whereas in other position the do not. Similarly, snares in position 8 shorten whereas in position 3 they do not. Once again, the context on the right and on the left is the same, there are both hi-hats. We exclude the interpretation of segmental influence. Nonetheless, it precedes position 9 where all hi-hats shortens; it is possible that there is an anticipatory effect of the final position that favors sound shortening in position 8.

4. Discussion

The study showed that there is indeed sound reduction, as well as the silences, when speeding up. The fact that silences shorten means gestures are closer to one another and may lead to articulatory overlap. The strategy here is to shorten gestures duration as well, to avoid such overlap in order to overcome the speeding up constraint and to maintain rhythmical pattern. However, we created a pattern that only contains coronals : [t' ts' t͡ʃi̥ ts' t't' ts' t͡ʃi̥ ts']. This pattern showed overlap between hi-hat and snares where hi-hat [ts'] were switching, though not completely, to [t͡ʃ]. We also found overlap between hi-hats and labials (e.g. [p'], [pf']) leading to [tsf].

5. Conclusion

Increasing beatboxing rate does not seem to affect production if both intergestural interval is reduced as well as gestures themselves. More participants are needed to study individual strategies and general tendencies. Human Beatboxing is an original, unexplored and rich paradigm to study the extent of the human vocal tract capacities and its biomechanics.

6. References

[1] Byrd, D., & Tan, C. C. (1996). Saying consonant clusters quickly. *Journal of Phonetics, 24(2)*, 263-282.

[2] Dehais Underdown, A., Vignes, P., Buchman, L. C., & Demolin, D. (2020, June). Rythme et contrôle articulatoire: étude préliminaire du Human Beatbox. In *Actes de la 6e conférence conjointe Journées d'Études sur la Parole (JEP, 31e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 1: Journées d'Études sur la Parole* (pp. 136-144). ATALA.

[3] Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The journal of the acoustical society of America, 101(6)*, 3728-3740.

[4] Lindblom, B. (1963). Spectroraphic Study of Vowel Reduction. *Journal of the acoustical society of America*, 1773-1781.

[5] Munhall, K., & Löfqvist, A. (1992). Gestural aggregation in speech: Laryngeal gestures. *Journal of Phonetics, 20(1)*, 111-126.

[6] Ostry, D. J., & Munhall, K. G. (1985). Control of rate and duration of speech movements. *The Journal of the Acoustical Society of America, 77(2)*, 640-648.